2.4 Expectation

Definition Let X be a discrete RV with possible values $x_1, \dots, x_n$. Then

$$E[X] := \sum_{i=1}^{n} x_i \, P[X = x_i]$$

is the *expected value* of X

Definition Let X be a continuous RV with density f. Then

$$E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

is the expected value of X (if the integral exists)

Example 36: A fire alarm exercise is announced for Monday between 2pm and 4pm. The alarm can go off any time between 2 and 4 pm and all times are equally likely. How long do we expect to wait?

Waiting time X : $0 \le X \le 2$ hrs

$$\text{Density } f(x) = \begin{cases} 1/2 & 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = \int_0^2 x \frac{1}{2}\, dx = \left[\frac{x^2}{4} \right]_0^2 = \frac{4}{4} = 1$$

2.2 Joint Distributions

Example 29. We have 9 batteries : 2 new, 3 partially charged, 4 empty. We randomly select 3 batteries out of the 9.

X number of new batteries selected

Y ——— " — partially charged batteries selected

Let $p(x,y) = P[X = x, Y = y]$. Values of $p(x,y)$

$$p(0,0) = \frac{\binom{4}{3}}{\binom{9}{3}} = \frac{4}{84} \qquad p(0,1) = \frac{\binom{3}{1}\binom{4}{2}}{\binom{9}{3}} = \frac{18}{84}$$

$$p(0,2) = \frac{\binom{3}{2}\binom{4}{1}}{\binom{9}{3}} = \frac{12}{84} \qquad p(0,3) = \frac{\binom{3}{3}}{\binom{9}{3}} = \frac{1}{84}$$

$$p(1,0) = \frac{\binom{2}{1}\binom{4}{2}}{\binom{9}{3}} = \frac{12}{84} \qquad p(1,1) = \frac{\binom{2}{1}\binom{3}{1}\binom{4}{1}}{\binom{9}{3}} = \frac{}{84}$$

$$p(1,2) = \frac{\binom{2}{1}\binom{3}{2}}{\binom{9}{3}} = \frac{6}{84}$$

$$p(2,0) = \frac{\binom{2}{2}\binom{4}{1}}{\binom{9}{3}} = \frac{4}{84} \qquad p(2,1) = \frac{\binom{2}{2}\binom{3}{1}}{\binom{9}{3}} = \frac{3}{84}$$

We have computed the joint probability mass function of X and Y

This can be summarized as follows (by multiples of $\frac{1}{84}$)

$X \backslash Y$	0	1	2	3	Sum
0	4	18	12	1	35
1	12	24	6	0	42
2	4	3	0	0	7
Sum	20	45	18	1	

probability of $X = i$

probability of $Y = i$

marginal probabilities

Generalize: Let X, Y be random variables (RVs).

$$F(x,y) := P[X \leq x, Y \leq y]$$

is the joint cumulative probability distribution of X and Y.

We get the distribution function of X from this as

$$f_X(x) = P[X \leq x] = P[X \leq x, Y < \infty]$$
$$= F(x, \infty)$$

The joint probability mass function of X and Y is

$$p(x,y) = P[X=x, Y=y]$$

for discrete X and Y. We get

$$P[X=x] = P\left(\bigcup_{j=1}^{n} [X=x, Y=y_j]\right)$$

$$= \sum_{j=1}^{n} P[X=x, Y=y_j] = \sum_{j=1}^{n} p(x,y_j)$$

Also: $P[Y=y] = \sum_{i=1}^{m} p(x_i,y)$

<u>Example 30</u> Assume the following statistical figures for a community of families

no	children	15 %
1	child	20 %
2	children	35 %
3	children	30 %

Probability of boys and girls is 50 % each

Calculate the joint mass function for boys and girls (i.e.: X = # boys, Y = # girls)

$P[X=0, Y=0] = P[\text{no children}] = 0.15$

$P[X=0, Y=1] = P[\text{"1 child"} \cap \text{"the one child is a girl"}]$

$= P[\text{"child is a girl"} \mid \text{"one child"}]\, P[\text{"one child"}]$

$= 0.5 \times 0.2 = 0.1$

$P[X=0, Y=2] = P[\text{"two children are girls"} \mid \text{"two children"}]$

$\times P[\text{"two children"}]$

$= (0.5)^2 \times 0.35 = 0.0875$

$P[X=0, Y=3] = P[\text{"three children are girls"} \mid \text{"three children"}]$

$\times P[\text{"three children"}]$

$= (0.5)^3 \times 0.3 = 0.0375$

$P[X=1, Y=1] = P[\text{"one boy, one girl"} \mid \text{"two children"}]$

$\times P[\text{"two children"}]$

$= 0.5 \times 0.35 = 0.175$

$P[X=1, Y=2] = P[\text{"one boy, two girls"} \mid \text{"3 children"}]$

$\times P[\text{"three children"}]$

$= 3 \cdot (0.5)^3 \times 0.3 = 3 \times 0.125 \times 0.3$

$= 0.375 \times 0.3 = 0.1125$

This can be translated into a table for the joint mass function

$x \backslash y$	0	1	2	3	sum
0	0.1500	0.1000	0.0875	0.0375	0.3750
1	0.100	0.1750	0.1125	0	0.3875
2	0.0875	0.1125	0	0	0.2000
3	0.0375	0	0	0	0.0375
sum	0.3750	0.3875	0.2000	0.0375	

<u>Definition 31</u> RVs X, Y are jointly continuous if there is $f: \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ s.th. for all ("reasonable") sets $C \subseteq \mathbb{R}\times\mathbb{R}$ we have

$$P[(X,Y)\in C] = \iint_{(x,y)\in C} f(x,y)\,dx\,dy$$

If $A, B \subseteq \mathbb{R}$, then

$$P[X\in A, Y\in B] = \int_B \int_A f(x,y)\,dx\,dy$$

The joint cumulative distribution of X, Y is

$$F(a,b) = P[X\le a, Y\le b] = \int_{-\infty}^{b}\int_{-\infty}^{a} f(x,y)\,dx\,dy$$

This implies

$$f(x,y) = \frac{\partial^2}{\partial x\,\partial y} F(x,y)$$ (whenever F is twice differentiable)

Note:

$$P[X=a,\ Y=b] = 0$$

$$P[X=a] = 0$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy \quad \text{is the}$$

(marginal) density of X

More generally:

$$P[X \in A] = P[X \in A,\ Y \in \mathbb{R}] = \int_A \int_{-\infty}^{\infty} f(x,y)\, dy\, dx$$

$$= \int_A f_X(x)\, dx$$

<u>Example</u> <u>32</u>: Let the joint density of X, Y be

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & x,y > 0 \\ 0 & \text{otherwise} \end{cases}$$

What are

$$P[X>1, Y<1],\quad P[X<a],\quad P[X<Y]\ ?$$

$$P[X>1, Y<1] = \int_0^1 \int_1^{\infty} 2e^{-x}e^{-2y}\, dx\, dy$$

$$= \int_0^1 2e^{-2y}\left(\int_1^{\infty} e^{-x}\, dx\right) dy$$

$$= \int_0^1 2e^{-2y}\left[-e^{-x}\right]_1^{\infty} dy = \int_0^1 2e^{-2y}(0+e^{-1})\, dy$$

$$= e^{-1}\int_0^1 2e^{-2y} dy = e^{-1}\left[-e^{-2y}\right]_0^1 = e^{-1}(1-e^{-2})$$